

온라인 쇼핑몰에서의 딥러닝 기반 패션 상품 추천 방식 연구

심현준, 성신웅
한국과학기술원 (KAIST)
flhy5836@kaist.ac.kr

Overview

- **딥러닝** 통해 **패션 상품 이미지의 유사도를 분석**하고,
이를 패션 기업의 **온라인 상품 추천 의사결정에 활용**
 - 방법론: 이미지 분석을 위한 딥러닝 기법 (Convolutional Neural Network, CNN)
및 상품 속성에 대한 분류 기법 (Support Vector Machine, SVM)
- **빅데이터 & 비즈니스 애널리틱스 활용연구**
 - 국내 패션 기업 K사와의 산학 협력 과제
 - 패션업계의 현업 운영 문제를 파악하고 구체적인 의사결정 지원을 통해 혁신 도모
 - **현업에서 실제 활용할 수 있는 의사결정 시스템 개발**
- **최신 기계학습 기법과 기업 데이터를 융합한 실제 사례 연구**

목차

1. 연구 배경
2. 연구 문제
3. 패션 상품 추천 방법론
4. 추천 결과 및 실제 적용 효과
5. 결론 및 향후 연구

1. 연구 배경

1. 연구 배경: 온라인 패션 시장의 성장

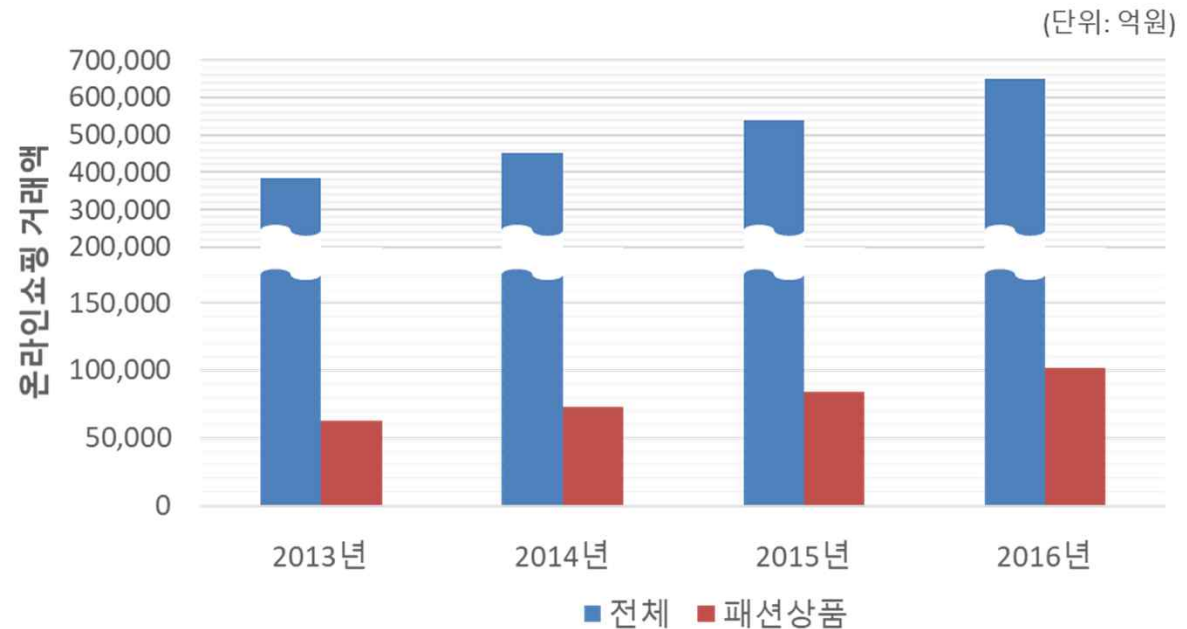
• 국내 온라인 패션 시장의 성장

- 국내 온라인 시장은 2014년 기준 매출액 세계 7위 규모^[1]
- 매년 두 자리의 높은 증가율을 보이며 2016년에는 약 65조 규모로 성장^[2]
- 이 중 의류 및 패션 관련 상품은 약 15.7% 으로 여행 서비스에 이어 두 번째로 큰 비중을 차지^[2]

Top 10 Countries, Ranked by Retail Ecommerce Sales Worldwide, 2013-2018
billions and % change

	2013	2014	2015	2016	2017	2018
1. China*	\$315.75	\$426.26	\$562.66	\$714.58	\$871.79	\$1,011.28
—% change	47.0%	35.0%	32.0%	27.0%	22.0%	16.0%
2. US**	\$264.28	\$305.65	\$349.06	\$394.43	\$442.55	\$493.89
—% change	16.5%	15.7%	14.2%	13.0%	12.2%	11.6%
3. UK**	\$70.39	\$82.00	\$93.89	\$104.22	\$114.64	\$124.96
—% change	17.0%	16.5%	14.5%	11.0%	10.0%	9.0%
4. Japan	\$62.13	\$70.83	\$79.33	\$88.06	\$96.87	\$106.07
—% change	17.9%	14.0%	12.0%	11.0%	10.0%	9.5%
5. Germany	\$51.91	\$63.38	\$73.46	\$82.93	\$91.97	\$99.33
—% change	21.7%	22.1%	15.9%	12.9%	10.9%	8.0%
6. France	\$34.21	\$38.36	\$42.62	\$46.13	\$49.71	\$53.26
—% change	13.2%	12.1%	11.1%	8.2%	7.8%	7.1%
7. South Korea	\$29.30	\$33.11	\$36.76	\$40.43	\$44.07	\$47.82
—% change	12.6%	13.0%	11.0%	10.0%	9.0%	8.5%
8. Canada	\$20.98	\$24.63	\$28.77	\$33.05	\$37.61	\$42.67
—% change	17.7%	17.4%	16.8%	14.9%	13.8%	13.5%
9. Russia	\$15.06	\$17.47	\$20.30	\$23.40	\$26.88	\$30.91
—% change	27.4%	16.0%	16.2%	15.3%	14.9%	15.0%
10. Brazil	\$13.34	\$16.28	\$18.80	\$21.34	\$23.79	\$26.17
—% change	28.0%	22.0%	15.5%	13.5%	11.5%	10.0%

Note: includes products or services ordered using the internet via any device, regardless of the method of payment or fulfillment; excludes travel and event tickets; *excludes Hong Kong; **forecast from Sep 2014
Source: eMarketer, Dec 2014
183295 www.eMarketer.com



^[1] Top 10 Countries, Ranked by Retail Ecommerce Sales Worldwide, 2013-2018, eMarketer, <http://www.emarketer.com/Chart/Top-10-Countries-Ranked-by-Retail-Ecommerce-Sales-Worldwide-2013-2018-billions-change/164369>

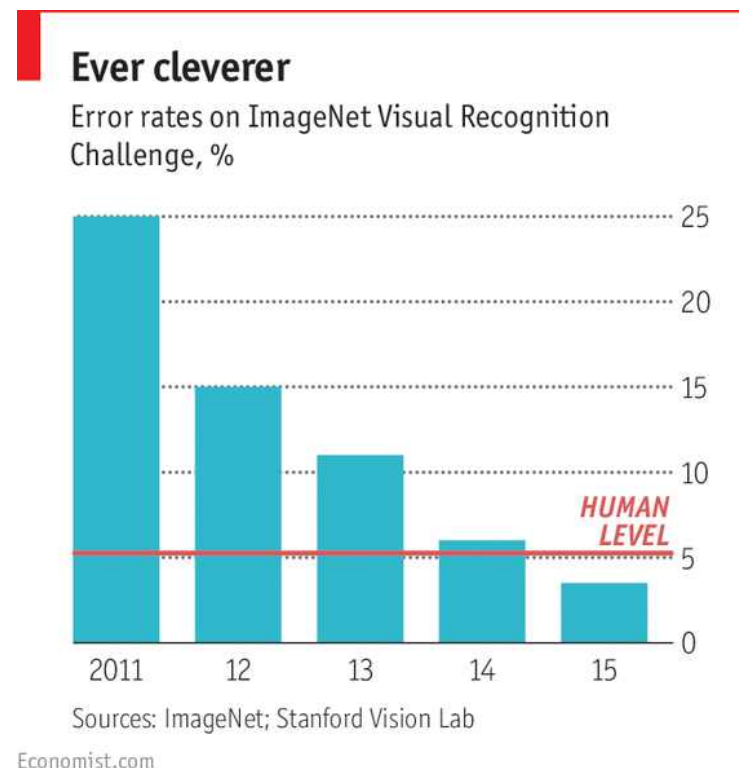
^[2] 2016년 12월 및 연간 온라인쇼핑 동향, 통계청

1. 연구 배경: 패션 상품 이미지 데이터의 증가와 딥러닝의 발달

- 패션 업계의 온라인 채널 강화 → 상품 이미지 데이터 증가
- 인공지능 및 딥러닝의 발달 → 이미지 인식 정확도 인간의 수준 능가

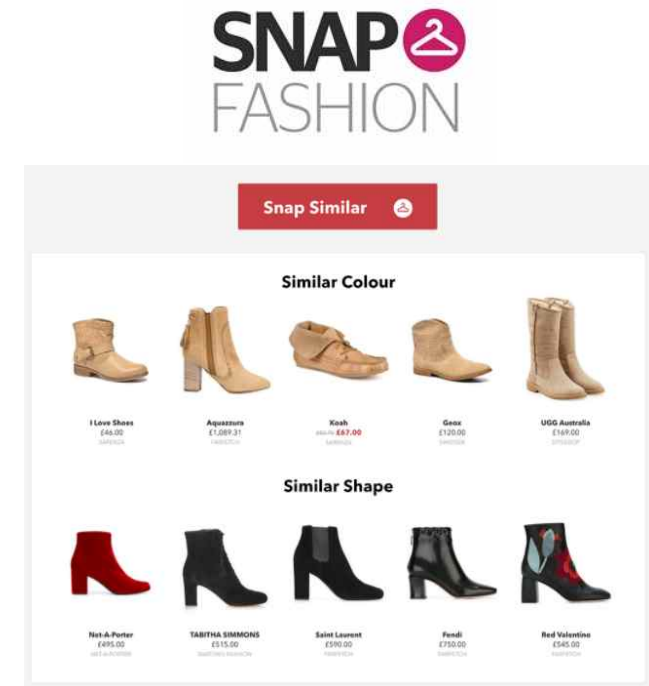
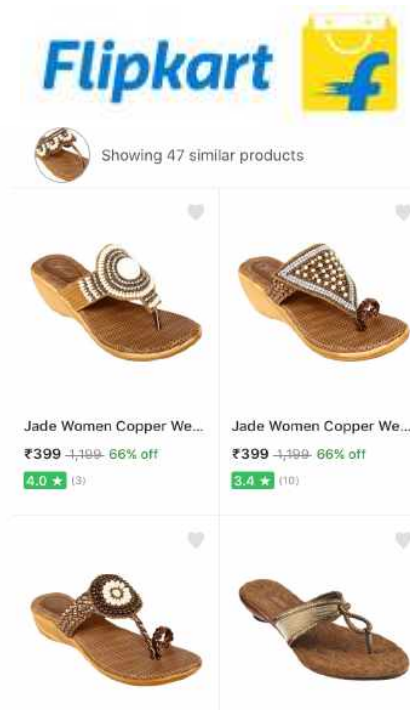
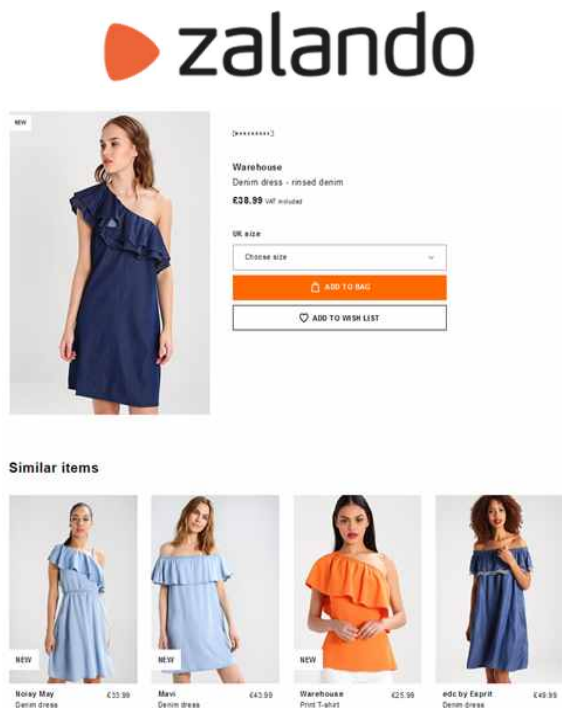


이미지 판별 딥러닝 모델 AlexNet의 결과 예시



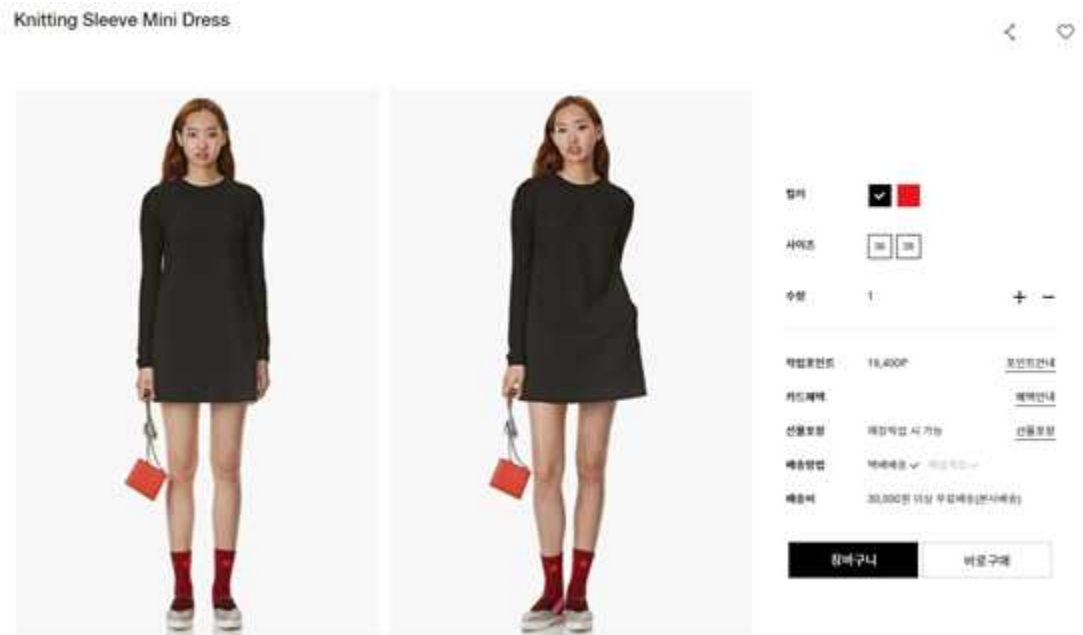
1. 연구 배경: 패션 상품 이미지 데이터의 증가와 딥러닝의 발달

- 패션 업계에서의 이미지 인식 인공지능 기술 활용
 - 유럽 최대 온라인 의류업체 Zalando
 - 인도 온라인 쇼핑몰 시장 1위 업체 Flipkart
 - 영국 온라인 패션 쇼핑몰 SNAP FASHION 등
- 상품 이미지를 활용한 인공지능 기술이 패션 온라인몰 경쟁력에 새로운 핵심



2. 연구 문제

2. 연구 문제: 온라인 쇼핑몰에서의 패션 상품 추천



다른 고객이 함께 본 상품



2. 연구 문제: 온라인 쇼핑몰에서의 패션 상품 추천

- “이 상품을 본 고객이 본 다른 상품” → **협업 필터링 (Collaborative filtering)**
 - 판매 및 클릭 데이터 기반으로 고객 간/상품 간 유사도 산출하여 상품 추천
 - 전통적으로 사용되어 왔고 다양한 상품군에 적용 가능
 - **Cold start issue**: 데이터가 충분히 쌓인 후에야 신뢰할 만한 추천 결과 보임
 - 패션 상품의 경우 계절성 및 유행으로 인해 상품 주기가 짧기 때문에 데이터가 충분히 쌓일 때 까지 기다려야 하는 방식은 적합하지 않음
 - **Long tail issue**: 상품 종류가 많을 경우 일부 상품의 데이터만 쌓이는 쓸림 현상 발생
 - 패션 상품은 매 시즌 수 백 종류 이상의 상품이 출시 – Long tail issue 가능성↑

2. 연구 문제: 온라인 쇼핑몰에서의 패션 상품 추천

- “이 상품을 본 고객이 본 다른 상품” → 협업 필터링 (Collaborative filtering)



2. 연구 문제: 온라인 쇼핑몰에서의 패션 상품 추천

- 콘텐츠 기반 필터링 (Content-based filtering)

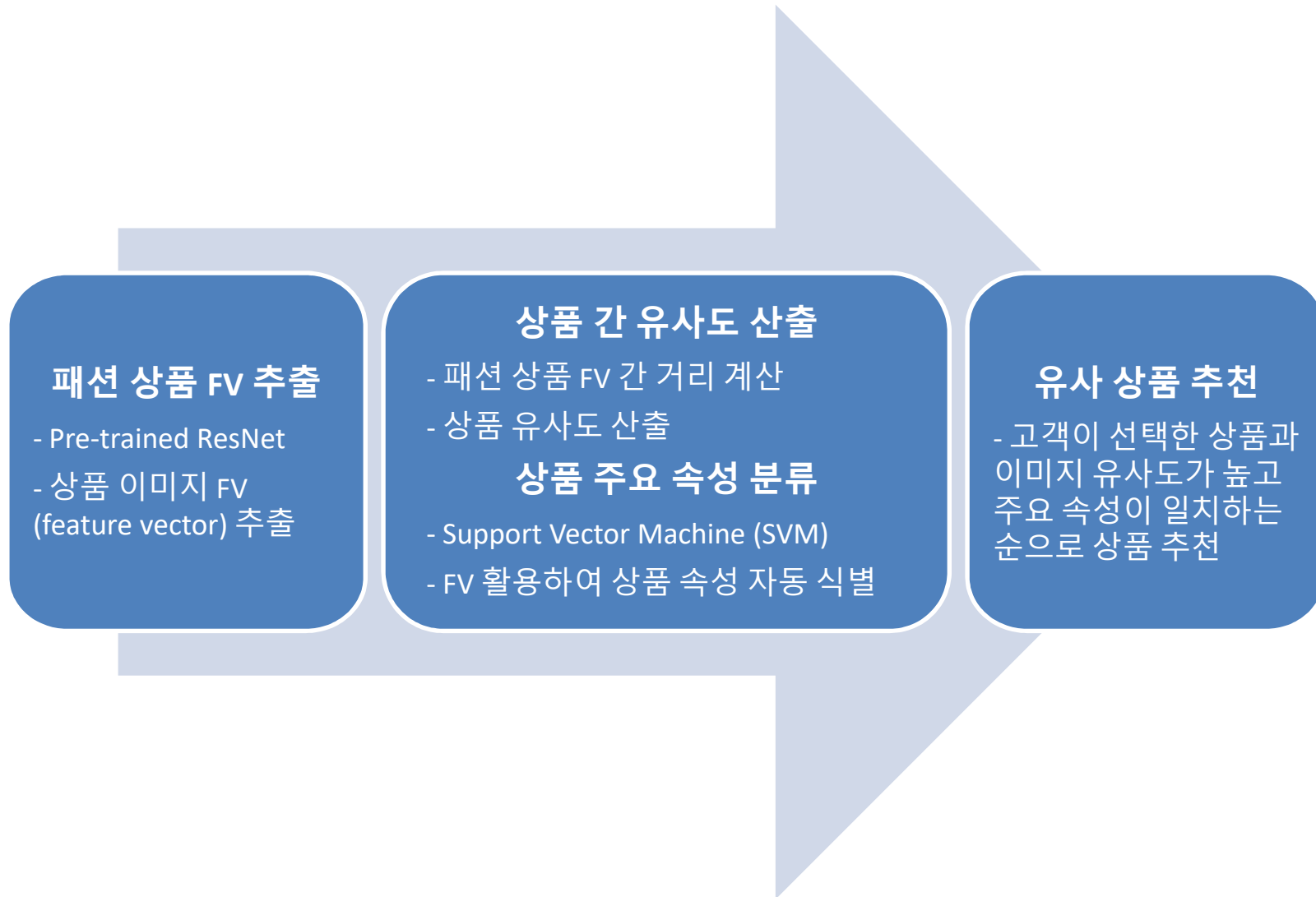
- 상품의 속성(e.g. 브랜드, 카테고리, 디자인 등)을 기반으로 상품 간 유사도 산출 & 고객이 선택한 상품과 유사한 상품 추천
 - 패션 상품 특성 상 시각적으로 유사한 상품 추천해 주면 구매로 이어질 확률 높음
 - **실 적용 이슈**
 - 대부분의 패션 기업에서 관리하는 상품 속성은 상위 레벨 데이터
 - 사람이 수기로 관리하여 속성에 대한 판단이 주관적
 - 속성만으로 상품의 시각적인 차이를 구분하기에 표현력이 충분하지 않음
- 현실적으로 잘 사용되지 않음

2. 연구 문제: 딥러닝 기술을 활용한 패션 상품 추천 연구

- 딥러닝 기술을 통해 패션 상품 이미지를 활용하여
상품의 시각적 정보를 효과적으로 추출 가능
→ 온라인 패션 쇼핑몰에서 고객이 선택한 상품과
시각적으로 유사한 상품을 추천하는 방식 제안
- 본 연구의 차별성
 - 딥러닝을 활용한 직관적인 패션 상품 추천 로직 제시
 - 딥러닝 결과와 함께 패션 상품의 주요 속성을 함께 고려
 - 국내 패션 기업의 자사 온라인 쇼핑몰에 제안 방식을 시범 적용하여
실제 효과 확인

3. 패션 상품 추천 방법론

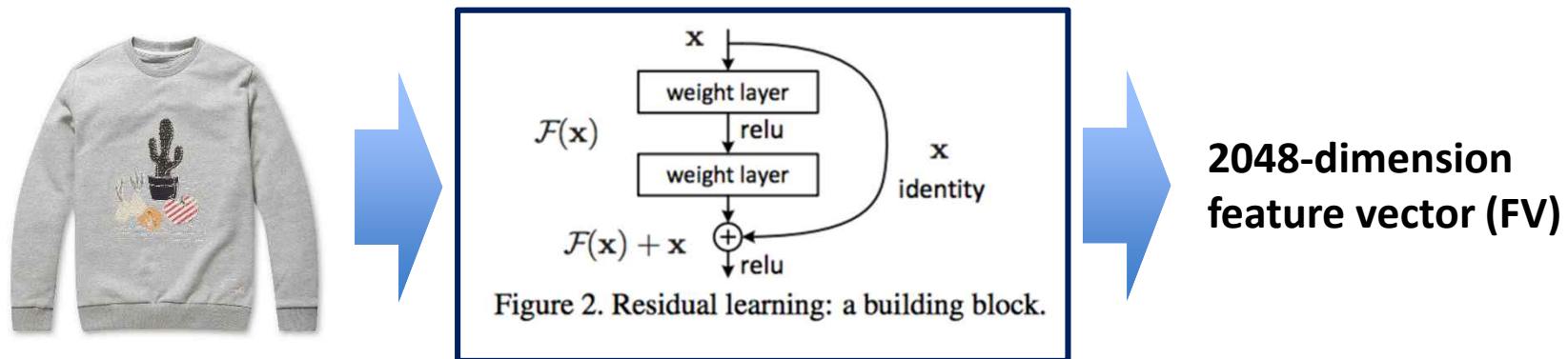
3. 방법론: 딥러닝 및 패션 상품 이미지를 활용한 상품 추천 프레임



3. 방법론: 딥러닝을 통한 상품 FV 추출 및 유사도 정량화

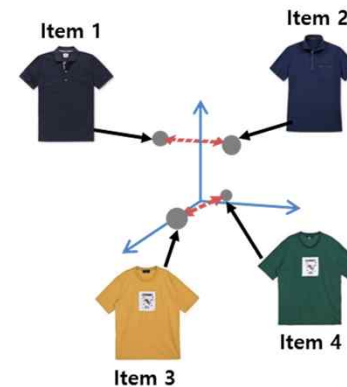
- CNN(Convolutional Neural Networks) 기반 이미지 프로세싱

- 오픈 소스로 활용되는 *Pre-trained ResNet** 활용
- 이미지가 갖는 추상적인 정보를 고차원의 벡터로 저장



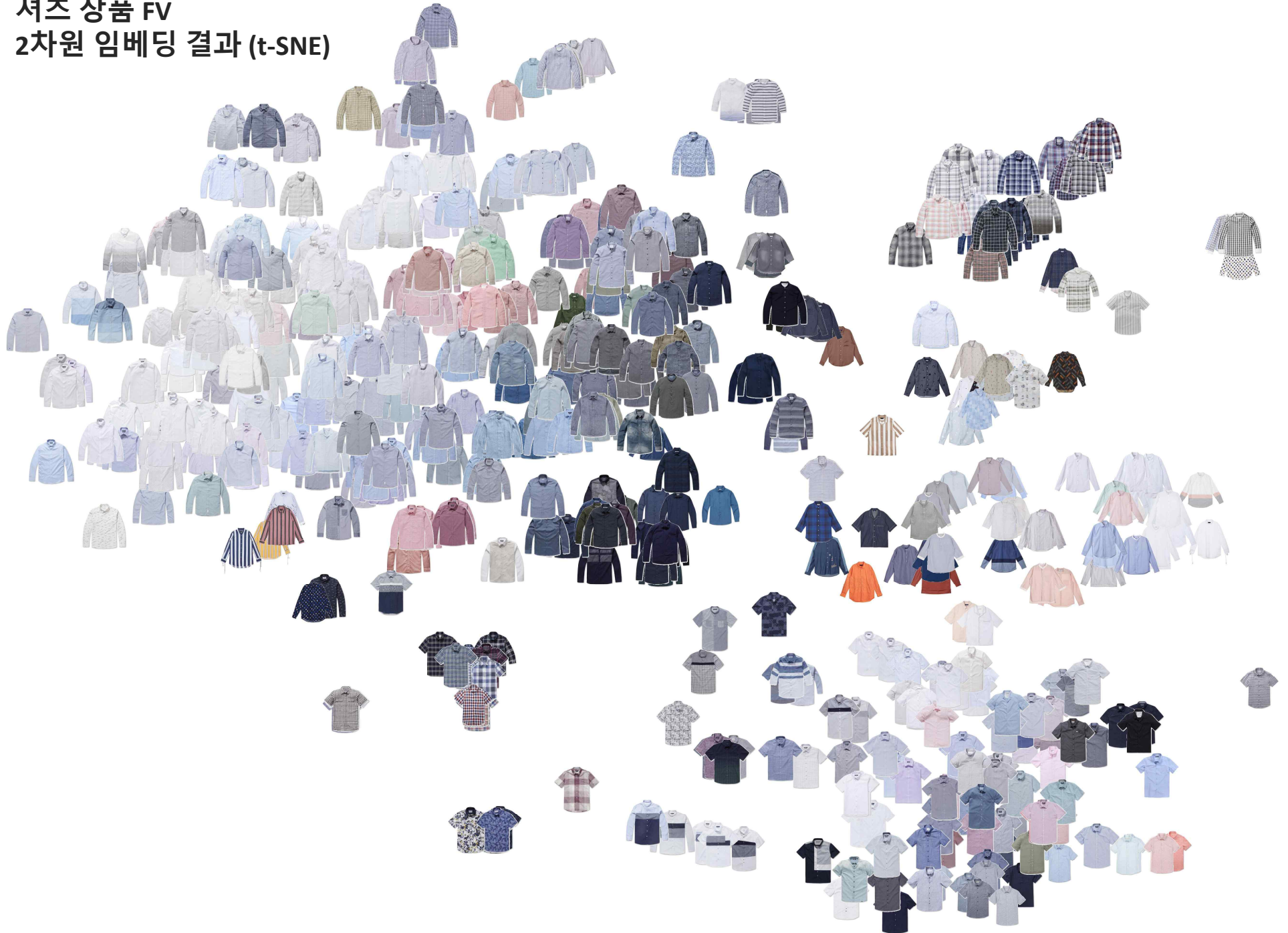
- 상품 간 유사도 정량화

- 각 상품들의 FV 사이 L2-distance를 계산
- 유사한 제품 간 거리는 작고,
유사하지 않은 제품 간 거리는 큼



*He, Kaiming, et al. "Deep residual learning for image recognition." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016.

셔츠 상품 FV
2차원 임베딩 결과 (t-SNE)



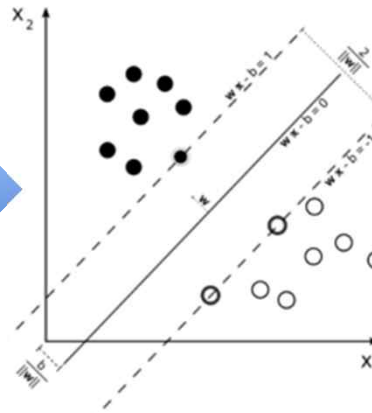
3. 방법론: 상품 주요 속성 자동 분류

- Support Vector Machine (SVM) 기법 기반 상품 속성 분류
 - 속성이 기 정의된 상품들의 FV를 학습/검증 데이터로 사용하여 속성이 정의되지 않은 상품들의 속성 분류



주요 속성 부위가
crop된 이미지의 FV

학습 및
검증



테스트



- 소매길이: 긴팔
- 넥형태: 라운드

	Training	Test
티셔츠 (T)		
소매길이	100%	100%
넥형태	100%	100%
스웨터 (W)		
소매길이	100%	96%
넥형태	99%	100%
여밈	100%	100%
셔츠 (S)		
소매길이	99%	100%
넥형태	96%	96%

	Training	Test
바지(P)		
기장	100%	100%
청바지	98%	97%
자켓 (J)		
점퍼풍	100%	100%
점퍼		
소매길이	100%	100%
밴드카라	98%	93%

3. 방법론: 딥러닝 결과와 상품 속성 정보를 결합한 패션 상품 추천 방식

1. 고객이 보고 있는 상품 (타깃 상품) 이미지의 FV와 거리가 가까운 k개의 상품 선별
2. 선별된 상품들의 속성이 타깃 상품의 속성과 비슷한 순으로 추천
 - 중요한 속성이 같은 상품 우선으로 추천
 - ex. 타깃 상품과 모든 속성 동일 → 소매 길이 동일 → 카라 유무 동일 ...



타깃 상품
- 반팔
- 카라

Step 1. 타깃 상품 이미지 FV와 거리가 가장 가까운 k=5개 상품 선별



Step 2. 타깃 상품 속성과 동일한 순으로 추천 (반팔 → 카라)



3. 방법론: 스타일링 상품 추천 방식

- 유사 상품 추천 알고리즘 + 스타일링 데이터
 - 브랜드 디자이너가 제안하는 서로 어울리는 제품 리스트 활용
 - 시각적으로 비슷한 상품들로 대체하여도 어울린다는 가정

Case 1	고객이 보고 있는 상품이 스타일링 리스트 안에 있는 경우
Case 1.1	스타일링 리스트 안에 한 세트(스타일) 있을 경우
추천 방법	해당 세트 상품들 추천 및 그 상품들과 유사한 N 세트 추천
Case 1.2	스타일링 리스트 안에 $1 < n \leq N$ 세트 있을 경우
추천 방법	해당 세트 상품들 추천 및 각 세트 상품들과 유사한 $(N - n + 1)/n$ 세트 추천
Case 1.3	스타일링 리스트 안에 $n > N$ 세트 있을 경우
추천 방법	해당 세트 상품들 모두 추천

Case 2	고객이 보고 있는 상품이 스타일링 리스트 안에 없는 경우
Case 2.1	해당 상품과 가장 유사한 동종 상품 k 개 추출하고, 그 중 스타일링 리스트에 속하는 상품이 k 개 있을 경우
추천 방법	해당 세트 상품들 추천 및 각 세트 상품들과 유사한 $(N - k + 1)/k$ 세트 추천
Case 2.2	해당 상품과 가장 유사한 동종 상품 k 개 추출하고, 그 중 스타일링 리스트에 속하는 상품이 없을 경우
추천 방법	스타일링 상품을 추천하지 않음



4. 추천 결과 및 실제 적용 효과

4. 추천 결과 및 실제 적용 효과

타깃 상품	추천 상품		
	1 	2 	3 
	4 	5 	6 

4. 추천 결과 및 실제 적용 효과

타깃 상품	추천 상품		
	1 	2 	3 
	4 	5 	6 

4. 추천 결과 및 실제 적용 효과

타깃 상품	추천 상품		
	1 	2 	3 
	4 	5 	6 

4. 추천 결과 및 실제 적용 효과

타깃 상품	추천 상품		
	1 	2 	3 
	4 	5 	6 

4. 추천 결과 및 실제 적용 효과

타킷 상품	추천 상품		
	1 	2 	3 
	4 	5 	6 

4. 추천 결과 및 실제 적용 효과

타깃 상품	추천 상품		
	1 	2 	3 
	4 	5 	6 

4. 추천 결과 및 실제 적용 효과

타킷 상품	추천 상품		
	1 	2 	3 
	4 	5 	6 

4. 추천 결과 및 실제 적용 효과

타킷 상품	추천 상품		
	1 	2 	3 
	4 	5 	6 

4. 추천 결과 및 실제 적용 효과

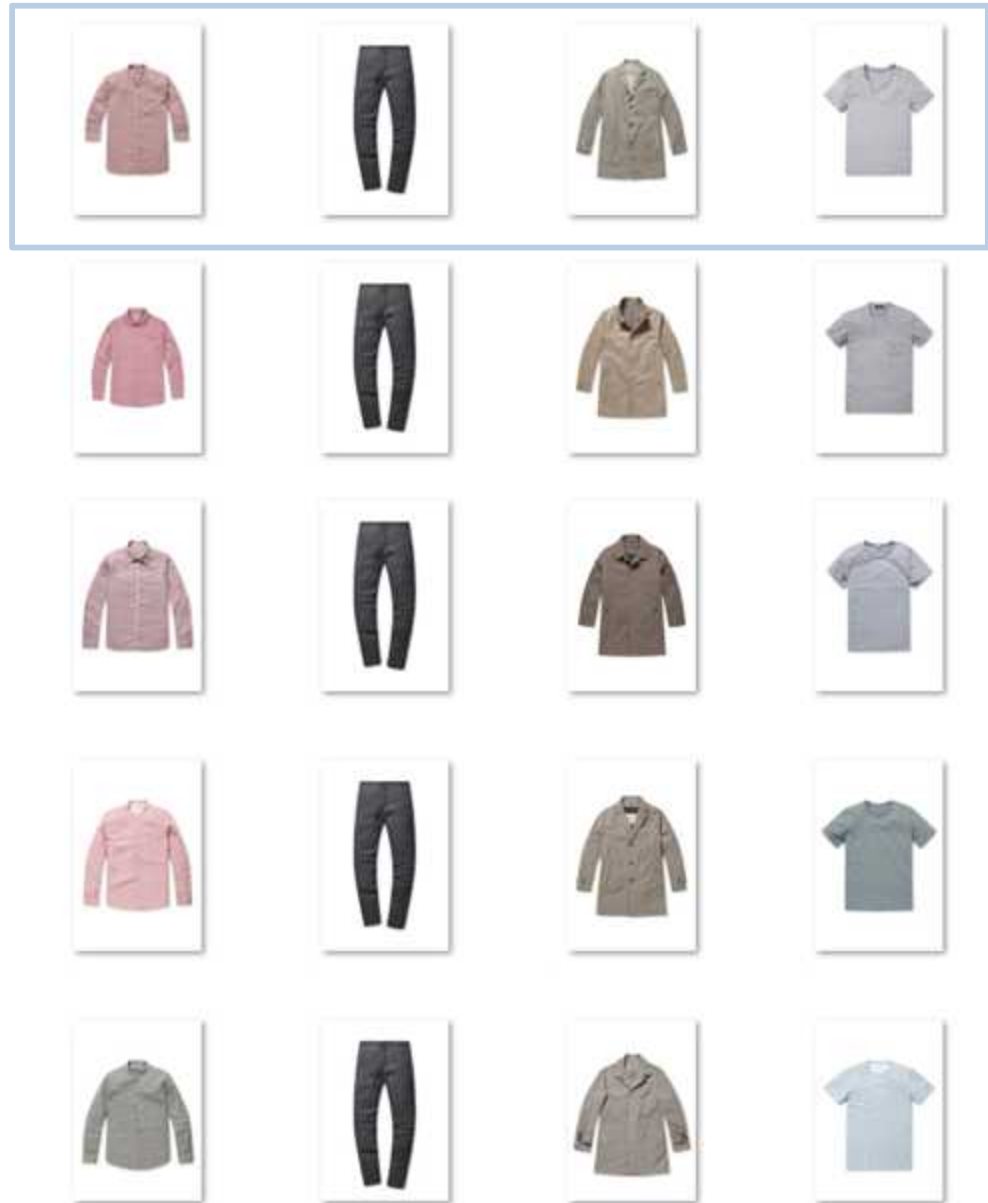
타킷 상품	추천 상품		
	1 	2 	3 
	4 	5 	6 

4. 추천 결과 및 실제 적용 효과

스타일링 상품 추천 결과 스타일북에 등록된 세트



타킷 상품

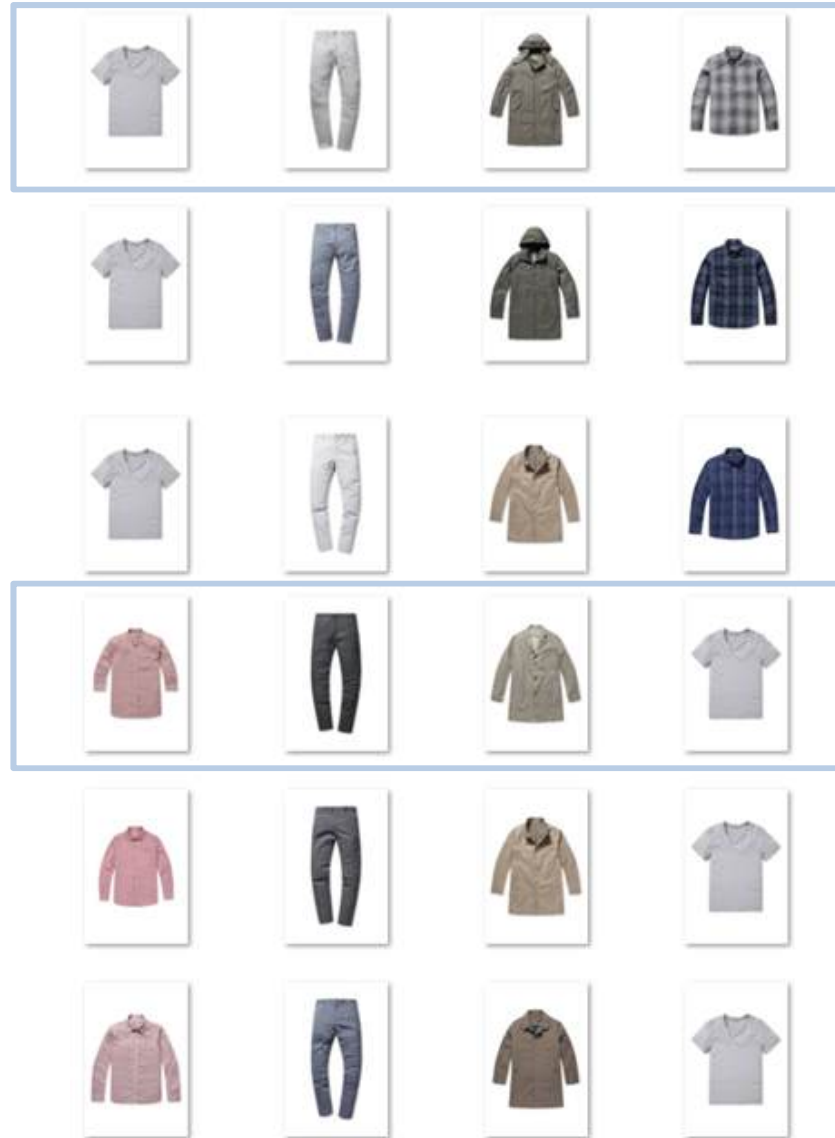


4. 추천 결과 및 실제 적용 효과

스타일링 상품 추천 결과



타킷 상품

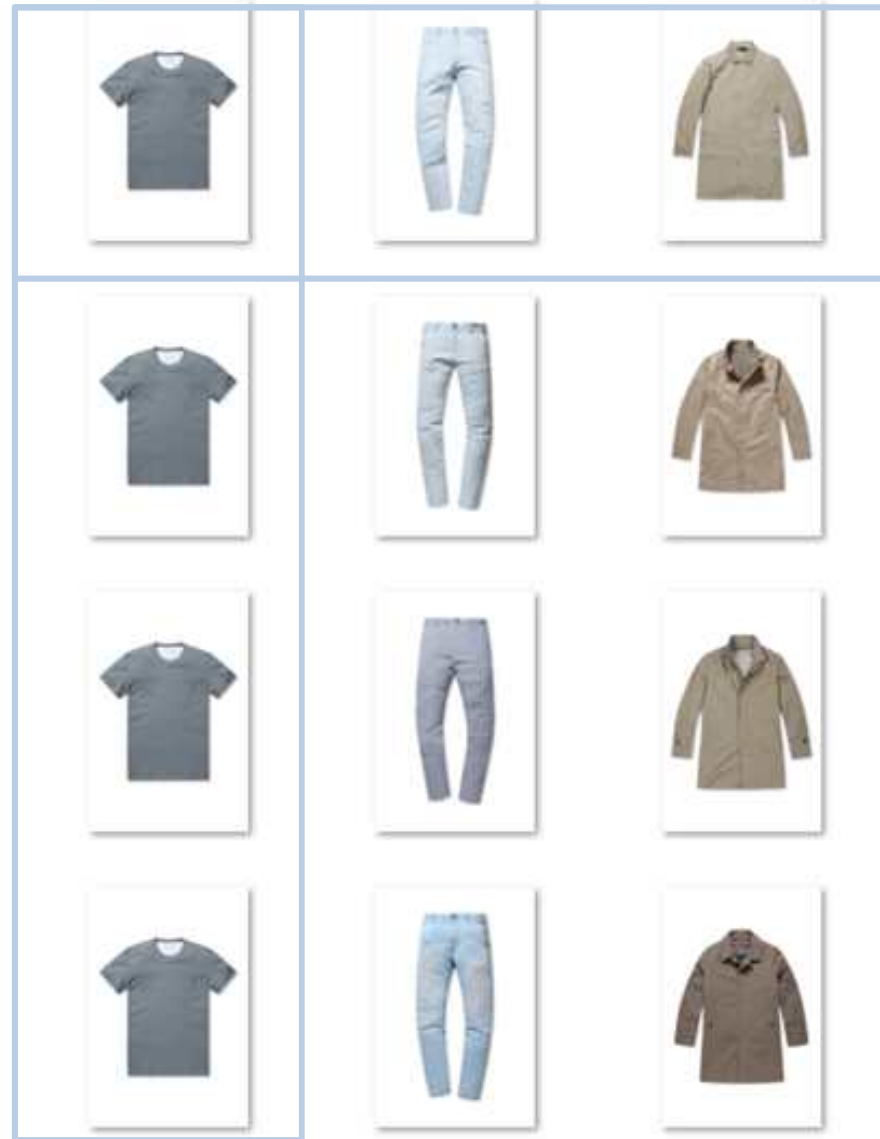


4. 추천 결과 및 실제 적용 효과

스타일링 상품 추천 결과



타깃 상품



4. 추천 결과 및 실제 적용 효과

- **국내 패션 기업 K사 온라인 쇼핑몰 시범 적용**

- 목적: 기존 추천 방식과의 비교를 통한 효과 검증
 - 기존 추천 방식: 협업 필터링 기반의 현업 자체 개발 방식
- 기간: 2017년 6월 5일 ~ 2017년 7월 2일 (4주)
- 대상 브랜드: 남성 캐주얼 7개 브랜드
- 추천 대상 Pool: 티 495개, 셔츠 572개, 바지 453개 → 총 1,520개 상품
- 타깃 상품: 3개 브랜드 각 품종 별 클릭 수 높은 상위 10개 상품
 - 클릭 수 내림차순으로 홀수 번째 상품을 실험군, 짝수 번째 상품을 대조군으로 선정

- **평가 척도**

- **추천 상품 방문자 비율**

= (타깃 상품 페이지에서 추천된 상품 방문자 수) / (타깃 상품 방문자 수) (%)

4. 추천 결과 및 실제 적용 효과

- 기존 방식과 제안 방식의 추천 상품 방문자 비율의 모평균의 차 통계 검정
 - 전체 표본에 대해, 제안 방식을 통한 추천 상품 방문자 비율 개선이 유의
 - 특히 바지 품종 및 브랜드 A 상품에 효과적
 - 바지의 경우 고객이 구매하고자 하는 상품이 시각적으로 명확한 편
 - 브랜드 A의 경우 상대적으로 더 실용적인 / 기본적인 디자인의 상품을 선보이는 브랜드

분석 표본		기존 방식 대비 제안 방식 (KAIST) 의 추천 상품 방문자 비율 증가량	귀무가설	P-value
전체		1.12배	$\mu_0 \geq \mu_{KAIST}$	0.175*
품종 별	티	1.10배	$\mu_0 \geq \mu_{KAIST}$	0.313
	바지	1.29배	$\mu_0 \geq \mu_{KAIST}$	0.133*
	셔츠	0.98배	$\mu_0 \leq \mu_{KAIST}$	0.448
브랜드 별	브랜드 A	1.73배	$\mu_0 \geq \mu_{KAIST}$	0.003***
	브랜드 B	1.08배	$\mu_0 \geq \mu_{KAIST}$	0.362
	브랜드 C	0.68배	$\mu_0 \leq \mu_{KAIST}$	0.016**

4. 추천 결과 및 실제 적용 효과

- 기존 방식과 제안 방식의 추천 상품 방문자 비율의 모평균의 차 통계 검정
 - 방문자 수가 적은 표본들에 대해서 제안 방식이 효과적 → Cold start issue 보완
 - 방문자 수 대비 판매량이 많은 표본들에 대해서 제안 방식이 효과적
 - 노출이 상대적으로 적은 상품에도 효과적 → Long tail issue 보완
 - 구매를 목적으로 상품을 클릭한 고객에게 제안 방식을 통한 추천 활용도 좋음

분석 표본	기존 방식 대비 제안 방식 (KAIST) 의 추천 상품 방문자 비율 증가량	귀무가설	P-value
방문자 수 하위 50%	1.57배	$\mu_0 \geq \mu_{KAIST}$	0.008***
방문자 수 대비 판매량 상위 50%	1.33배	$\mu_0 \geq \mu_{KAIST}$	0.041**

5. 결론 및 향후 연구

5. 결론 및 향후 연구

- 대표적인 기계학습 방법인 딥러닝을 활용하여, 온라인 패션몰에서 고객이 선택한 상품과 시각적으로 유사한 상품을 추천하는 방식 제안
- 이론적으로 발전된 최신 기법을 실제 현장에 적용한 사례
- 본 연구에서 제안한 추천 방식
→ 기업 내부적으로 실무에 활용될 수 있도록 추가적인 분석 작업 진행 중
- Future research
 - 시각적 유사도 뿐만 아니라 클릭/판매 데이터 함께 고려한 추천 방식 개발
 - 고객이 온라인에서 패션 상품 쇼핑 시 중요하게 생각하는 상품의 주요 속성 검증